

On the Method of Least Squares. By J. W. L. Glaisher, M.A.,
F.R.S.

The present paper contains certain notes on the method of least squares; as they are to some extent connected with my previous paper "On the Solution of the Equations in the Method of Least Squares" (*Monthly Notices*, vol. xxxiv. pp. 311-334), the sections are numbered as if this were a continuation of that paper.

§ 14. The results of the treatment of a system of linear equations by the method of least squares may be stated as follows:—

Let

$$\begin{aligned} a_1x + b_1y + c_1z + \dots + f_1t &= n_1, \\ a_2x + b_2y + c_2z + \dots + f_2t &= n_2, \\ \dots &\dots \dots \dots \dots \dots \\ a_mx + b_my + c_mz + \dots + f_mt &= n_m, \end{aligned}$$

be the m equations of condition, x, y, z, \dots, t being the μ unknown quantities, and n_1, n_2, \dots, n_m the m quantities obtained by observation.* The latter are assumed to be equally good—i.e. they are assumed to be such that the mean error of each of them is equal to the mean error ε of a standard observation.

The normal equations giving x_0, y_0, \dots, t_0 , the "most probable" values of x, y, \dots, t are

$$\begin{aligned} (aa)x_0 + (ab)y_0 + (ac)z_0 + \dots + (af)t_0 &= (an), \\ (ba)x_0 + (bb)y_0 + (bc)z_0 + \dots + (bf)t_0 &= (bn), \\ \dots &\dots \dots \dots \dots \dots \\ (fa)x_0 + (fb)y_0 + (fc)z_0 + \dots + (ff)t_0 &= (fn), \end{aligned}$$

where, for example,

$$\begin{aligned} (aa) &\text{ denotes } a_1^2 + a_2^2 + \dots + a_m^2, \\ (ab) &\text{, } \quad a_1b_1 + a_2b_2 + \dots + a_mb_m, \end{aligned}$$

so that $(ab) = (ba)$, &c.

* The notation, as in the previous paper, is derived from the case of six unknowns, x, y, z, w, u, t , the coefficients being denoted by a, b, c, d, e, f , with suffixes attached. All the formulæ in the paper, however, have reference to the general case of μ unknowns, x, y, z, \dots, t , with corresponding coefficients a, b, c, \dots, f .

The signs of the quantities n_1, n_2, \dots, n_m have been changed, so that the equations of condition are written

$$a_1x + b_1y + c_1z + \dots + f_1t = n_1, \quad \&c.$$

instead of

$$a_1x + b_1y + c_1z + \dots + f_1t + n_1 = 0, \quad \&c.,$$

as it is somewhat more convenient to have the quantities determined by observation on the right hand side of the equations. This change is also made in Oppolzer's *Lehrbuch zur Bahnbestimmung der Kometen und Planeten*, vol. ii. p. 313 (1880). In my previous paper Gauss's and Encke's form (i.e. the second of the above forms) was retained.

The values of x_0, y_0, \dots are therefore

$$x_0 = \frac{\begin{vmatrix} (an), (ab), (ac), \dots (af') \\ (bn), (bb), (bc), \dots (bf') \\ \dots \dots \dots \dots \dots \\ (fn), (fb), (fc), \dots (ff') \end{vmatrix}}{\begin{vmatrix} (aa), (ab), (ac), \dots (af') \\ (ba), (bb), (bc), \dots (bf') \\ \dots \dots \dots \dots \dots \\ (fa), (fb), (fc), \dots (ff') \end{vmatrix}},$$

$$y_0 = \frac{\begin{vmatrix} (aa), (an), (ac), \dots (af') \\ (ba), (bn), (bc), \dots (bf') \\ \dots \dots \dots \dots \dots \\ (fa), (fn), (fc), \dots (ff') \end{vmatrix}}{\begin{vmatrix} (aa), (ab), (ac), \dots (af') \\ (ba), (bb), (bc), \dots (bf') \\ \dots \dots \dots \dots \dots \\ (fa), (fb), (fc), \dots (ff') \end{vmatrix}},$$

$$z_0 = \&c.$$

The weights p_x, p_y, \dots of x_0, y_0, \dots are

$$p_x = \frac{\begin{vmatrix} (aa), (ab), \dots (af') \\ (ba), (bb), \dots (bf') \\ \dots \dots \dots \dots \dots \\ (fa), (fb), \dots (ff') \end{vmatrix}}{\begin{vmatrix} (bb), (bc), \dots (bf') \\ (cb), (cc), \dots (cf') \\ \dots \dots \dots \dots \dots \\ (fb), (fc), \dots (ff') \end{vmatrix}},$$

$$p_y = \frac{\begin{vmatrix} (aa), (ab), \dots (af') \\ (ba), (bb), \dots (bf') \\ \dots \dots \dots \dots \dots \\ (fa), (fb), \dots (ff') \end{vmatrix}}{\begin{vmatrix} (aa), (ac), \dots (af') \\ (ca), (cc), \dots (cf') \\ \dots \dots \dots \dots \dots \\ (fa), (fc), \dots (ff') \end{vmatrix}},$$

$$p_z = \&c.;$$

and the sum of the squares of the residuals, viz. (vv) ,

$$= \frac{\begin{vmatrix} (aa), (ab), \dots (af), (an) \\ (ba), (bb), \dots (bf), (bn) \\ \dots \dots \dots \dots \dots \\ (fa), (fb), \dots (ff), (fn) \\ (na), (nb), \dots (nf), (nn) \end{vmatrix}}{\begin{vmatrix} (aa), (ab), \dots (af) \\ (ba), (bb), \dots (bf) \\ \dots \dots \dots \dots \\ (fa), (fb), \dots (ff) \end{vmatrix}}.$$

If therefore $\epsilon_x, \epsilon_y, \dots$ denote the mean errors of x_0, y_0, \dots

$$\epsilon^{2x} = \frac{1}{m-\mu} \frac{\begin{vmatrix} (bb), (bc), \dots (bf) \\ (cb), (cc), \dots (cf) \\ \dots \dots \dots \dots \\ (fb), (fc), \dots (ff) \end{vmatrix}}{\begin{vmatrix} (aa), (ab), \dots (af) \\ (ba), (bb), \dots (bf), (bn) \\ \dots \dots \dots \dots \\ (fa), (fb), \dots (ff), (fn) \\ (na), (nb), \dots (nf), (nn) \end{vmatrix}},$$

$$\frac{\begin{vmatrix} (aa), (ab), \dots (af) \\ (ba), (bb), \dots (bf) \\ \dots \dots \dots \dots \\ (fa), (fb), \dots (ff) \end{vmatrix}^2}{\begin{vmatrix} (aa), (ab), \dots (af) \\ (ba), (bb), \dots (bf), (bn) \\ \dots \dots \dots \dots \\ (fa), (fb), \dots (ff), (fn) \\ (na), (nb), \dots (nf), (nn) \end{vmatrix}},$$

$$\epsilon_y^2 = \frac{1}{m-\mu} \frac{\begin{vmatrix} (aa), (ac), \dots (af) \\ (ca), (cc), \dots (cf) \\ \dots \dots \dots \dots \\ (fa), (fc), \dots (ff) \end{vmatrix}}{\begin{vmatrix} (aa), (ab), \dots (af), (an) \\ (ba), (bb), \dots (bf), (bn) \\ \dots \dots \dots \dots \\ (fa), (fb), \dots (ff), (fn) \\ (na), (nb), \dots (nf), (nn) \end{vmatrix}},$$

$$\frac{\begin{vmatrix} (aa), (ab), \dots (af) \\ (ba), (bb), \dots (bf) \\ \dots \dots \dots \dots \\ (fa), (fb), \dots (ff) \end{vmatrix}^2}{\begin{vmatrix} (aa), (ab), \dots (af), (an) \\ (ba), (bb), \dots (bf), (bn) \\ \dots \dots \dots \dots \\ (fa), (fb), \dots (ff), (fn) \\ (na), (nb), \dots (nf), (nn) \end{vmatrix}},$$

$$\epsilon_z^2 = \&c.;$$

where μ is the number of unknowns.

The form of the expressions for $\epsilon_x^2, \epsilon_y^2, \dots$ is remarkable; the numerators consist of the product of two determinants, of which one contains $\mu-1$, and the other $\mu+1$ rows, and the denominator is the square of a determinant of μ rows.

Denoting the last-mentioned determinant, viz.

$$\begin{vmatrix} (aa), (ab), \dots (af) \\ (ba), (bb), \dots (bf) \\ \dots \dots \dots \\ (fa), (fb), \dots (ff) \end{vmatrix}$$

by ∇ , then, in the numerator of the expression for ε_x^2 , the first determinant is formed from ∇ by omitting the first line and column, and the second determinant is formed from ∇ by bordering it with the elements $(an), (bn), \dots (nn), (nf), \dots (na)$; similarly, the first determinant in the numerator of ε_y^2 is derived from ∇ by omitting the second line and second column, and so on. The second determinant and the denominator, ∇^2 , are the same in all the expressions.

The foregoing values of x_0, y_0, \dots as determinants and of p_x, p_y, \dots as quotients of determinants may of course be written down at once from the equations which determine them, and no formal proof is needed.* The value of (vv) is taken from § 7 of the previous paper, and a direct proof of this result is given in the next section (§ 15). In § 6 expressions were found for the auxiliaries $(bb.1), (bc.1), (cc.2), \dots$, as quotients of determinants.

§ 15. The method by which it was shown in § 7 that (vv) was equal to

$$\frac{1}{\nabla} \begin{vmatrix} (aa), (ab), \dots (af), (an) \\ (ba), (bb), \dots (bf), (bn) \\ \dots \dots \dots \dots \\ (fa), (fb), \dots (ff), (fn) \\ (na), (nb), \dots (nf), (nn) \end{vmatrix},$$

depended upon the known equality of (vv) and $(nn.6)$. The expression for (vv) may, however, be very simply established, without assuming this equality, as follows:—

$$\begin{aligned} (vv) &= \Sigma(a_1x_0 + b_1y_0 + c_1z_0 + \dots - n_1)^2, \\ &= (aa)x_0^2 + (bb)y_0^2 + \dots + 2(ab)x_0y_0 + 2(ac)x_0z_0 + \dots \\ &\quad - 2(an)x_0 - 2(bn)y_0 - \dots + (nn), \\ &= x_0\{(aa)x_0 + (ab)y_0 + \dots + (af)z_0 - (an)\} \\ &\quad + y_0\{(ba)x_0 + (bb)y_0 + \dots + (bf)z_0 - (bn)\} \\ &\quad + \dots \dots \dots \dots \dots \dots \\ &\quad + t_0\{(fa)x_0 + (fb)y_0 + \dots + (ff)z_0 - (fn)\} \\ &\quad - (an)x_0 - (bn)y_0 - \dots - (fn)t_0 + (nn). \end{aligned}$$

* In a paper entitled "Over het gebruik van determinanten bij de methode der kleinste kwadraten," printed in the *Nieuw Archief voor Wiskunde*, Deel I. pp. 179-188 (1875), Mr. Van Geer gives the determinant values of x_0, y_0, \dots and of their weights. The determinants which form the denominators of the latter are, through some inadvertence, erroneous, although the equations from which they are derived are correctly stated.

In virtue of the normal equations which x_0, y_0, \dots satisfy, each line in this expression vanishes except the last, and therefore

$$\begin{aligned}
 (vv) &= -(an)x_0 - (bn)y_0 - \dots - (fn)t_0 + (nn), \\
 &= -\frac{(an)}{\nabla} \begin{vmatrix} (an), (ab), \dots (af) \\ (bn), (bb), \dots (bf) \\ \dots \\ (fn), (fb), \dots (ff) \end{vmatrix} - \frac{(bn)}{\nabla} \begin{vmatrix} (aa), (an), \dots (af) \\ (ba), (bn), \dots (bf) \\ \dots \\ (fa), (fn), \dots (ff) \end{vmatrix} - \dots + (nn), \\
 &= (-)^n \frac{1}{\nabla} \begin{vmatrix} (an), (bn), \dots (fn), (nn) \\ (aa), (ab), \dots (af), (an) \\ \dots \\ (ba), (bb), \dots (bf), (bn) \\ (fa), (fb), \dots (ff), (fn) \end{vmatrix}, \\
 &= \frac{1}{\nabla} \begin{vmatrix} (aa), (ab), \dots (af), (an) \\ (ba), (bb), \dots (bf), (bn) \\ \dots \\ (fa), (fb), \dots (ff), (fn) \\ (na), (nb), \dots (nf), (nn) \end{vmatrix}.
 \end{aligned}$$

It will be noticed that the analysis employed in this section merely amounts to the use of the following almost obvious theorem:—taking three letters only for simplicity, if x, y, z be given by the equations

$$\begin{aligned}
 \alpha_1 x + \beta_1 y + \gamma_1 z &= h_1, \\
 \alpha_2 x + \beta_2 y + \gamma_2 z &= h_2, \\
 \alpha_3 x + \beta_3 y + \gamma_3 z &= h_3,
 \end{aligned}$$

then

$$px + qy + rz - k = \frac{\begin{vmatrix} p, q, r, k \\ \alpha_1, \beta_1, \gamma_1, h_1 \\ \alpha_2, \beta_2, \gamma_2, h_2 \\ \alpha_3, \beta_3, \gamma_3, h_3 \end{vmatrix}}{\begin{vmatrix} \alpha_1, \beta_1, \gamma_1 \\ \alpha_2, \beta_2, \gamma_2 \\ \alpha_3, \beta_3, \gamma_3 \end{vmatrix}}.$$

If the signs of n_1, n_2, \dots, n_m be changed, the signs of the expressions x_0, y_0, \dots are changed, but the values of $(vv), p_x, p_y, \dots, \epsilon_x^2, \epsilon_y^2, \dots$ remain unaltered.

§16. *Theorem:* The determinant ∇ , viz.

$$\begin{array}{l|l} (aa), (ab), \dots (af) & \dots \dots \dots (1) \\ (ba), (bb), \dots (bf) & \\ \dots \dots \dots & \\ (fa), (fb), \dots (ff) & \end{array}$$

is equal to the sum of the squares of p determinants, p denoting the number of combinations of m things μ together.

To prove this, write at length the determinant ∇ ,

$$\begin{array}{l|l} a_1^2 + a_2^2 \dots + a_m^2, \quad b_1 a_1 \dots + b_m a_m, \quad \dots, \quad f_1 a_1 \dots + f_m a_m & \\ a_1 b_1 + a_2 b_2 \dots + a_m b_m, \quad b_1^2 \dots + b_m^2, \quad \dots, \quad f_1 b_1 \dots + f_m b_m & \\ \dots \dots \dots & \\ a_1 f_1 + a_2 f_2 \dots + a_m f_m, \quad b_1 f_1 \dots + b_m f_m, \quad \dots, \quad f_1^2 \dots + f_m^2 & \end{array}$$

and from the $[m]^\mu$ determinants of which this is the sum select that which is formed from the first terms of the constituents of the first column, the second terms of the second column, \dots and the μ -th terms of the last column: this determinant

$$= a_1 b_2 \dots f_\mu \begin{array}{l|l} a_1, a_2, \dots a_\mu & \\ b_1, b_2, \dots b_\mu & \\ \dots \dots \dots & \\ f_1, f_2, \dots f_\mu & \end{array} .$$

Similarly the determinant formed from the second terms of the first column, the first terms of the second column, and the same constituents as before in the case of the other columns, is

$$a_2 b_1 c_3 \dots f_\mu \begin{array}{l|l} a_2, a_1, a_3, \dots a_\mu & \\ b_2, b_1, b_3, \dots b_\mu & \\ \dots \dots \dots & \\ f_2, f_1, f_3, \dots f_\mu & \end{array} ,$$

which is

$$= -a_2 b_1 c_3 \dots f_\mu \begin{array}{l|l} a_1, a_2, \dots a_\mu & \\ b_1, b_2, \dots b_\mu & \\ \dots \dots \dots & \\ f_1, f_2, \dots f_\mu & \end{array} ,$$

and in this way it is easily seen that the coefficient of the determinant

$$\begin{array}{l|l} a_1, a_2, \dots a_\mu & \dots \dots \dots (2) \\ b_1, b_2, \dots b_\mu & \\ \dots \dots \dots & \\ f_1, f_2, \dots f_\mu & \end{array}$$

is $\Sigma \pm a_1 b_2 c_3 \dots f_\mu$, the summation referring to all permutations of the suffixes 1, 2, ..., μ : but this quantity is itself equal to the determinant (2), so that the resulting expression is equal to the square of (2), and, interchanging the lines and columns in the determinants of the form (2), we have the result

$$\left| \begin{array}{c} (aa), (ab), \dots (af) \\ (ba), (bb), \dots (bf) \\ \dots \dots \dots \\ (fa), (fb), \dots (ff) \end{array} \right| = \Sigma \left| \begin{array}{c} a_1, b_1, \dots f_1 \\ a_2, b_2, \dots f_2 \\ \dots \dots \dots \\ a_\mu, b_\mu, \dots f_\mu \end{array} \right|^2 \dots \dots \dots \quad (3)$$

where the summation symbol refers to the suffixes which in the different determinants are the $[m]^\mu \div [m]^\mu$ sets of μ numbers that can be formed from the m numbers 1, 2, ..., m .

Using a recognised notation, the determinant (2) may be conveniently written

$$(a_1 b_2 \dots f_\mu),$$

and the theorem is that the determinant (1) is equal to $\Sigma (a_1 b_2 \dots f_\mu)^2$.

§17. It can be easily seen, in the same manner, that the determinant

$$\left| \begin{array}{c} (an), (ab), (ac), \dots (af) \\ (bn), (bb), (bc), \dots (bf) \\ \dots \dots \dots \\ (fn), (fb), (fc), \dots (ff) \end{array} \right| \dots \dots \dots \quad (4)$$

is equal to

$$\Sigma \left| \begin{array}{c} a_1, b_1, \dots f_1 \\ a_2, b_2, \dots f_2 \\ \dots \dots \dots \\ a_\mu, b_\mu, \dots f_\mu \end{array} \right| \left| \begin{array}{c} n_1, b_1, \dots f_1 \\ n_2, b_2, \dots f_2 \\ \dots \dots \dots \\ n_\mu, b_\mu, \dots f_\mu \end{array} \right|$$

or, as this expression may be written,

$$\Sigma (a_1 b_2 \dots f_\mu) (n_1 b_2 \dots f_\mu).$$

the Σ having the same meaning as at the end of the last section, and the number of terms—i.e. of products of pairs of determinants—being as before $= [m]^\mu \div [\mu]^\mu$.

§18. The values of x_0, y_0, \dots are therefore

$$x_0 = \frac{\Sigma (a_1 b_2 \dots f_\mu) (n_1 b_2 \dots f_\mu)}{\Sigma (a_1 b_2 \dots f_\mu)^2},$$

$$y_0 = \frac{\Sigma (a_1 b_2 \dots f_\mu) (a_1 n_2 \dots f_\mu)}{\Sigma (a_1 b_2 \dots f_\mu)^2},$$

$$z_0 = \&c.;$$

and it thus appears that the values of x, y, \dots given by the method of least squares are in fact those obtained by combining linearly in the manner indicated the values found by solving each set of μ equations which can be formed from the m given equations. Expressing this more in detail, we have the following rule:—

From the m equations of condition we can form p sets of μ equations, p denoting for brevity $[m]^\mu \div [\mu]^\mu$. Solve each of these sets of equations by the determinant method, and let the resulting values be

$$\text{for } x, \frac{\alpha_1}{\lambda_1}, \frac{\alpha_2}{\lambda_2}, \dots, \frac{\alpha_p}{\lambda_p},$$

$$\text{for } y, \frac{\beta_1}{\lambda_1}, \frac{\beta_2}{\lambda_2}, \dots, \frac{\beta_p}{\lambda_p},$$

&c.,

where $\alpha_1, \dots, \beta_1, \dots, \lambda_1, \dots$ are the actual determinants which occur in the solutions (*i.e.* so that any factors common to both numerator and denominator in any of the fractions are not to be thrown out); then

$$x_0 = \frac{\lambda_1 \alpha_1 + \lambda_2 \alpha_2 + \dots + \lambda_p \alpha_p}{\lambda_1^2 + \lambda_2^2 + \dots + \lambda_p^2},$$

$$y_0 = \frac{\lambda_1 \beta_1 + \lambda_2 \beta_2 + \dots + \lambda_p \beta_p}{\lambda_1^2 + \lambda_2^2 + \dots + \lambda_p^2},$$

$z_0 = \text{&c.}$

The rule may also be stated in a slightly different manner thus:— solve each set of equations and let the system of values be

$$\text{for } x, A_1, A_2, \dots, A_p,$$

$$\text{for } y, B_1, B_2, \dots, B_p,$$

&c.;

then

$$x_0 = \frac{\lambda_1^2 A_1 + \lambda_2^2 A_2 + \dots + \lambda_p^2 A_p}{\lambda_1^2 + \lambda_2^2 + \dots + \lambda_p^2},$$

$$y_0 = \frac{\lambda_1^2 B_1 + \lambda_2^2 B_2 + \dots + \lambda_p^2 B_p}{\lambda_1^2 + \lambda_2^2 + \dots + \lambda_p^2},$$

$z_0 = \text{&c.}$

where $\lambda_1, \lambda_2, \dots, \lambda_p$ are as before: that is, they are the determinants whose constituents are the coefficients which occur on the left-hand side of the different sets of equations.

§19. To illustrate the formulæ, take the simple case of $m=4, \mu=3$.

By §16,

$$\left| \begin{array}{l} (aa), (ab), (ac) \\ (ba), (bb), (bc) \\ (ca), (cb), (cc) \end{array} \right| = (a_1 b_2 c_3)^2 + (a_1 b_2 c_4)^2 + (a_1 b_3 c_4)^2 + (a_2 b_3 c_4)^2,$$

by §17,

$$\left| \begin{array}{l} (an), (ab), (ac) \\ (bn), (bb), (bc) \\ (cn), (cb), (cc) \end{array} \right| = (a_1 b_2 c_3)(n_1 b_2 c_3) + (a_1 b_2 c_4)(n_1 b_2 c_4) \\ + (a_1 b_3 c_4)(n_1 b_3 c_4) + (a_2 b_3 c_4)(n_2 b_3 c_4),$$

and by §18,

$$x_0 = \frac{(a_1 b_2 c_3)(n_1 b_2 c_3) + (a_1 b_2 c_4)(n_1 b_2 c_4) + (a_1 b_3 c_4)(n_1 b_3 c_4) + (a_2 b_3 c_4)(n_2 b_3 c_4)}{(a_1 b_2 c_3)^2 + (a_1 b_2 c_4)^2 + (a_1 b_3 c_4)^2 + (a_2 b_3 c_4)^2},$$

$$y_0 = \frac{(a_1 b_2 c_3)(a_1 n_2 c_3) + (a_1 b_2 c_4)(a_1 n_2 c_4) + (a_1 b_3 c_4)(a_1 n_3 c_4) + (a_2 b_3 c_4)(a_2 n_3 c_4)}{(a_1 b_2 c_3)^2 + (a_1 b_2 c_4)^2 + (a_1 b_3 c_4)^2 + (a_2 b_3 c_4)^2},$$

$$z_0 = \frac{(a_1 b_2 c_3)(a_1 b_2 n_3) + (a_1 b_2 c_4)(a_1 b_2 n_4) + (a_1 b_3 c_4)(a_1 b_3 n_4) + (a_2 b_3 c_4)(a_2 b_3 n_4)}{(a_1 b_2 c_3)^2 + (a_1 b_2 c_4)^2 + (a_1 b_3 c_4)^2 + (a_2 b_3 c_4)^2}.$$

§20. As a numerical example of the process, I now consider the system of four equations which Gauss himself employed to illustrate the method of least squares, and which has generally been adopted as the standard example by writers on the subject. These equations are

$$\left. \begin{array}{l} x - y + 2z = 3 \\ 3x + 2y - 5z = 5 \\ 4x + y + 4z = 21 \\ -x + 3y + 3z = 14 \end{array} \right\} \begin{array}{l} (i) \\ (ii) \\ (iii) \\ (iv) \end{array};$$

and the normal equations derived from them are

$$\left. \begin{array}{l} 27x_0 + 6y_0 = 88 \\ 6x_0 + 15y_0 + z_0 = 70 \\ y_0 + 54z_0 = 107 \end{array} \right\},$$

which give

$$x_0 = \frac{49154}{19899}, \quad y_0 = \frac{2617}{737}, \quad z_0 = \frac{12707}{6633}.$$

Taking the first three equations (i), (ii), (iii), the values of x, y, z , given by them are

$$x = \frac{P_1}{P}, \quad y = \frac{P_2}{P}, \quad z = \frac{P_3}{P},$$

where

$$P = \begin{vmatrix} 1, & -1, & 2 \\ 3, & 2, & -5 \\ 4, & 1, & 4 \end{vmatrix} = 35,$$

$$P_1 = \begin{vmatrix} 3, & -1, & 2 \\ 5, & 2, & -5 \\ 21, & 1, & 4 \end{vmatrix} = 90,$$

$$P_2 = \begin{vmatrix} 1, & 3, & 2 \\ 3, & 5, & -5 \\ 4, & 21, & 4 \end{vmatrix} = 115,$$

$$P_3 = \begin{vmatrix} 1, & -1, & 3 \\ 3, & 2, & 5 \\ 4, & 1, & 21 \end{vmatrix} = 65.$$

Similarly solving equations (i), (ii), (iv),

$$x = \frac{Q_1}{Q}, \quad y = \frac{Q_2}{Q}, \quad z = \frac{Q_3}{Q},$$

where the values of the determinants are

$$Q = \begin{vmatrix} 1, & -1, & 2 \\ 3, & 2, & -5 \\ -1, & 3, & 3 \end{vmatrix} = 47,$$

$$Q_1 = 122, \quad Q_2 = 167, \quad Q_3 = 93.$$

Solving the systems (i), (iii), (iv), and (ii), (iii), (iv), and denoting the values of x, y, z , by

$$\frac{R_1}{R}, \quad \frac{R_2}{R}, \quad \frac{R_3}{R}$$

and

$$\frac{S_1}{S}, \quad \frac{S_2}{S}, \quad \frac{S_3}{S},$$

the values of the determinants $R, R_1, \dots, S, S_1, \dots$ are found to be

$$R = 33, \quad R_1 = 78, \quad R_2 = 113, \quad R_3 = 67 \\ S = -124, \quad S_1 = -304, \quad S_2 = -444, \quad S_3 = -236.$$

The unreduced values obtained from the four sets of equations are thus

$$\text{for } x, \frac{90}{35}, \frac{122}{47}, \frac{78}{33}, \frac{304}{124},$$

$$\text{for } y, \frac{115}{35}, \frac{167}{47}, \frac{113}{33}, \frac{444}{124},$$

$$\text{for } z, \frac{65}{35}, \frac{93}{47}, \frac{67}{33}, \frac{236}{124},$$

and therefore

$$x_0 = \frac{90 \times 35 + 122 \times 47 + 78 \times 33 + 304 \times 124}{(35)^2 + (47)^2 + (33)^2 + (124)^2},$$

$$y_0 = \frac{115 \times 35 + 167 \times 47 + 113 \times 33 + 444 \times 124}{(35)^2 + (47)^2 + (33)^2 + (124)^2},$$

$$z_0 = \frac{65 \times 35 + 93 \times 47 + 67 \times 33 + 236 \times 124}{(35)^2 + (47)^2 + (33)^2 + (124)^2},$$

giving

$$x_0 = \frac{49154}{19899}, \quad y_0 = \frac{70659}{19899}, \quad z_0 = \frac{38121}{19899},$$

which agree with the values given by the normal equations.

In general, if the values obtained by solving the equations be given in any form such as, e.g., vulgar fractions in their lowest terms, so that the values of x are

$$\frac{18}{7}, \quad \frac{122}{47}, \quad \frac{26}{11}, \quad \frac{76}{31},$$

then these quantities are to have the respective weights

$$P^2, \quad Q^2, \quad R^2, \quad S^2,$$

viz.

$$x_0 = \frac{\frac{18}{7} \times (35)^2 + \frac{122}{47} \times (47)^2 + \frac{26}{11} \times (33)^2 + \frac{76}{31} \times (124)^2}{(35)^2 + (47)^2 + (33)^2 + (124)^2},$$

and the weights are the same in the case of the corresponding values of y and z .

§21. In the absence of any method such as that of least squares, if we had to determine the best values of x, y, \dots from a system of m linear equations, it would be natural to first solve every set of μ equations which could be formed from the m equations, and to compare the different values of x, y, \dots thus obtained. The question to be decided would be how to combine the different values, and perhaps the method which would first suggest itself would be to take the arithmetic mean of the values found for x as the adopted value of x , and similarly in the case of

y, z, \dots ; it is clear, however, that this mode of treatment would not be satisfactory, as certain of the sets of equations would be better suited for the accurate determination of x, y, \dots than others.

It appears from § 18 and 22 that in proceeding according to the method of least squares we assign to each system of values of x, y, \dots a weight proportional to the square of the determinant whose constituents are the coefficients in the set of μ equations from which the values are derived.

It is evident that this determinant affords a good measure of the precision with which x, y, \dots are determined by the set of equations, representing as it does the common denominator in the values of these quantities. In the method of least squares the square of the determinant is taken as the weight, and in consequence the sign of the determinant is immaterial, and the common denominator of x_0, y_0, \dots , being a sum of squares, is always positive.

§ 22. The determinant which forms the numerator of (vv) , viz.

$$\begin{vmatrix} (aa), (ab), \dots (af), (an) \\ (ba), (bb), \dots (bf), (bn) \\ \dots \dots \dots \dots \\ (fa), (fb), \dots (ff), (fn) \\ (na), (nb), \dots (nf), (nn) \end{vmatrix} \quad (\mu + 1 \text{ rows}) \dots \dots \dots (5)$$

is of the same form as ∇ and only differs from it by including the letters n : the determinant (5) is therefore

$$\equiv \nabla \begin{vmatrix} a_1, b_1, \dots f_1, n_1, \\ a_2, b_2, \dots f_2, n_2, \\ \dots \dots \dots \dots \dots \\ a_\mu, b_\mu, \dots f_\mu, n_\mu \\ a_{\mu+1}, b_{\mu+1}, \dots f_{\mu+1}, n_{\mu+1} \end{vmatrix}^2 \dots \dots \dots (6)$$

i.e. it is equal to the sum of the squares of all the determinants whose constituents are the coefficients and right-hand members in the different sets of $(\mu + 1)$ equations which can be formed from the m equations; the number of such determinants is $[m]^{\mu+1} \div [\mu + 1]^{\mu+1}$. It thus appears that (vv) cannot vanish unless every one of these determinants is equal to zero—i.e. unless the m equations are all consistent with one another, and equivalent to only μ independent equations. If $(vv) = 0$, the mean (or probable) errors of x_0, y_0, \dots also vanish; and this is as it should be, for it is clear that the mean (or probable) errors of x_0, y_0, \dots can only be zero when the m equations determine x, y, \dots uniquely—i.e. are equivalent to only μ independent equations.

§ 23. If the number of equations exceeds the number of un-

knowns by unity—i.e. if $m=\mu+1$ —then the expression (6) consists of only a single determinant, so that (vv) is a complete square, and therefore $\sqrt{(vv)}$ is a linear function of $n_1, n_2, \dots, n_{\mu-1}$, its value being

$$\frac{1}{\nabla} \left| \begin{array}{ccccc} a_1, & b_1, & \dots & f_1, & n_1 \\ a_2, & b_2, & \dots & f_2, & n_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{\mu}, & b_{\mu}, & \dots & f_{\mu}, & n_{\mu} \\ a_{\mu+1}, & b_{\mu+1}, & \dots & f_{\mu+1}, & n_{\mu+1} \end{array} \right|$$

If in this determinant capital letters be used to denote the minors of the corresponding italic letters, then

$$\sqrt{(vv)} = \frac{n_1 N_1 + n_2 N_2 + \dots + n_{\mu+1} N_{\mu+1}}{\sqrt{(N_1^2 + N_2^2 + \dots + N_{\mu+1}^2)}},$$

$$x_0 = -\frac{A_1 N_1 + A_2 N_2 + \dots + A_{\mu+1} N_{\mu+1}}{N_1^2 + N_2^2 + \dots + N_{\mu+1}^2},$$

$$y_0 = -\frac{B_1 N_1 + B_2 N_2 + \dots + B_{\mu+1} N_{\mu+1}}{N_1^2 + N_2^2 + \dots + N_{\mu+1}^2},$$

$$z_0 = \&c.$$

In the case of the foregoing example

$$\begin{aligned} \sqrt{(vv)} &= \frac{1}{\sqrt{\nabla}} \left| \begin{array}{cccc} 1, & -1, & 2, & 3 \\ 3, & 2, & -5, & 5 \\ 4, & 1, & 4, & 21 \\ -1, & 3, & 3, & 14 \end{array} \right| \\ &= \frac{14P - 21Q + 5R - 3S}{\sqrt{(P^2 + Q^2 + R^2 + S^2)}} \\ &= \frac{14 \times 35 - 21 \times 47 + 5 \times 33 + 3 \times 124}{\sqrt{(35^2 + 47^2 + 33^2 + 124^2)}} \\ &= \frac{40}{\sqrt{19899}}. \end{aligned}$$

§ 24. The determinants which form the denominators of p_x, p_y, \dots are also sums of squares of determinants; for example, the denominator of p_x , viz.:

the number of determinants being $[m]^{\mu-1} \div [\mu-1]^{\mu-1}$.

Thus, in the example, the denominator of p_x

$$\begin{aligned}
 &= \left| \begin{array}{cc} -1, & 2^{-2} + \\ 2, & -5 \end{array} \right| + \left| \begin{array}{cc} -1, & 2^{-2} + \\ 1, & 4 \end{array} \right| + \left| \begin{array}{cc} -1, & 2^{-2} + \\ 3, & 3 \end{array} \right|^2 \\
 &+ \left| \begin{array}{cc} 2, & -5 \\ 1, & 4 \end{array} \right|^2 + \left| \begin{array}{cc} 2, & -5 \\ 3, & 3 \end{array} \right|^2 + \left| \begin{array}{cc} 1, & 4 \\ 3, & 3 \end{array} \right|^2 \\
 &= 1 + 36 + 81 + 169 + 441 + 81 \\
 &= 809,
 \end{aligned}$$

and therefore

$$p_x = \frac{19899}{809}.$$

It may be observed that if the determinant which forms the denominator in any of the expressions for the weights is equal to zero, so also is the numerator, for if each of the determinants in (8) vanishes, then each of the determinants which form the right-hand side of (3) vanishes also.

§ 25. Since all the determinants involved in the statement of results in § 14 are equal to sums of squares, it follows that they can never be negative. In § 5 it was shown that all the products $(aa)(bb.1)$, $(aa)(bb.1)(cc.2)$, &c., are of the form (1), and it is therefore evident that the auxiliaries $(bb.1)$, $(cc.2)$, &c., cannot be negative. A shorter proof of this fact was, however, given in § 7.

§ 26. The case, $\mu=2$, of the theorem in § 16 is a very well known result; but the only place I know of in which the general theorem is enunciated occurs in an investigation by Professor C. Niven of the vibrations of a dynamical system where the particles are subject to small frictional forces, printed in the "Cambridge Senate-House Problems and Riders for 1878," pp. 188-191. The proof given in § 16 is the same as Professor Niven's.

X X

Professor Cayley, to whom I communicated the results of §§ 16-18 some time ago was, I found, already acquainted with them; but it seems possible that they may not have been published before, as it would not be easy to express them without the aid of determinants, and the only paper I have met with in which an explicit use of determinants is made in connection with the solution of equations in the method of least squares is that of Mr. Van Geer referred to in the note in § 14.

On the Possible Performance of an Object-Glass for Star-Gazing.
By Edward Sang, Esq.

It having been proposed to compute the curvatures for an object-glass, with a view to obtaining the least possible aberration in the image of a star, the preliminary question arose as to how the computations should be conducted.

In the preparation of formulæ for the amount of spherical aberration, the sines and cosines of arcs are represented by two or three terms of the series which truly express them, and therefore such formulæ are only approximative, and the results obtained by them are to be regarded as guides to more accurate determinations. Our ultimate resort is to trace strictly the course of each pencil of light. When the thicknesses are taken into account, the application of the formulæ becomes as laborious as the direct trigonometrical calculation itself; wherefore it was determined to follow the trigonometrical method throughout.

The computation thus takes the form of a series of trials applicable only to the particular case in hand, and we have so to arrange these trials as to make them exhaustive, and so also as to throw light on analogous cases.

The proposition as it occurs in practice is this:—"Given two discs of glass, to construct of them an object-glass which shall give the best possible result." In the present instance that result is to be the formation of the image of an exceedingly minute luminous object; the correlative matters of the flatness of the field of view and of the performance towards the edge of that field not being taken into account. Now, the refractions by the two kinds of glass are data in the problem and fix the amount of the secondary chromatic aberration; wherefore our enquiry must be mainly directed to that part of the total error, which depends on the sphericity of the surfaces—that being the only matter under the control of the constructor.

The case actually proposed was to make an aplanatic combination from two discs, one of *hard-crown*, the other of *dense flint* glass, having an aperture of 7.5 inches, with a thickness in the rough of .75; the indices of refraction being given in Chance's list as under, and the focal distance to be about 100 inches.